Abstract

A database comprising information on amino acid sequences of proteins with one or more known biological functions, and further comprising information on importance scores with regard to the appearance of said biological functions added for each amino acid residue constituting said amino acid sequences, and a method of preparing an alignment of a protein stored in the database and a polypeptide with unknown biological function which comprises the steps of calculating a homology score to the coincidence of each constituent amino acid under consideration of the importance score for the appearance of a biological function, and preparing an alignment representing homology of sites where said importance is high.